Optimizing Neural Networks for Activity Recognition in Daily Living

A Case Study Using Signal Processing and Smartwatch Sensors

Klemens Waldhör
FOM University
Nuremberg, Germany
Email: klemens.waldhoer@fom.de

Philipp Müller
Business Information Systems, FOM University
Nuremberg, Germany
Email: muellerphilipp17.08@gmail.com

Abstract— **This study explores the impact of various signal processing techniques on neural network performance for activity recognition using smartwatch sensor data. Four common Activities of Daily Living (ADLs) including drinking, tumbling, teeth brushing, and walking, are evaluated. Signal processing methods, Gaussian filtering, Principal Component Analysis (PCA), Fourier Transform (FT), Empirical Mode Decomposition (EMD), and Hilbert-Huang Transform (HHT), are systematically assessed for their effectiveness in improving neural network classification accuracy. Multiple deep learning architectures, including Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), Gated Recurrent Units (GRU), and Convolutional Neural Networks (CNN), are implemented and compared. Results reveal that signal processing techniques significantly enhance the performance of RNN models, whereas other architectures (LSTM, GRU, CNN) achieve high accuracy (>99%) without additional signal preprocessing. Additionally, a hybrid CNN-LSTM model was successfully deployed on a Samsung Galaxy Watch 6, to classify ADLs within a smartwatch. However, practical implementation challenges, such as battery consumption and the necessity for on-device learning capabilities, are identified. This research provides valuable insights into optimizing neural network performance for wearable computing in resource-constrained environments.**

Keywords— Activity Recognition; Signal Processing; Neural Networks; Wearable Computing; Smartwatch Sensors.

I. INTRODUCTION

Germany is undergoing a pronounced demographic transition marked by an increasingly elderly population and persistently low birth rates [1]. By 2049, estimates suggest Germany will require between 280 000 and 690 000 additional care professionals to meet the needs of its aging citizens [2]. To bridge this gap, healthcare systems must turn to technological innovations that streamline patient monitoring and support clinical decision-making.

In this context, wearable devices, most notably smartwatches, have shown considerable promise. Equipped with accelerometers, gyroscopes and heart-rate sensors, they offer continuous, non-invasive tracking of Activities of Daily Living (ADLs), potentially enhancing diagnostic anamnesis and enabling rapid emergency response to events such as falls or acute cardiac episodes [3]-[5].

Yet, deploying advanced neural-network models directly on smartwatches introduces significant challenges: limited processing power, constrained memory, and the need to preserve battery life [6]-[9]. Effective real-time classification of complex movements therefore hinges on balancing model accuracy with resource efficiency.

This study investigates whether neural networks trained on raw smartwatch sensor data can accurately distinguish between a wide range of human movements, whether incorporating signal-processing techniques such as Fourier or wavelet transforms can boost classification performance, and how different time-series encoding methods affect the classification accuracy of these models in multi-class activity recognition.

The rest of the paper is structured as follows. Section II presents the related work. Section III describes the methodology, and Section IV the results. We conclude the work in Section V.

II. RELATED WORK

Time series classification represents one of ten challenging problems in data mining research [10]. The noise in time series data poses a particular challenge that requires sophisticated approaches to address effectively. Previous research by Waldhör and Lutze has successfully demonstrated the real-time recognition of drinking activities using smartwatches [11][12], establishing the feasibility of ADL detection in wearable devices.

The development of RNNs can be traced back to the early 1980s with Hopfield networks [13], designed as content-addressable memory systems. Significant progress was achieved in the 1990s with the introduction of fully connected RNN architectures by researchers like Jeffrey Elman and Michael I. Jordan [14]. However, these networks struggled with the vanishing gradient problem, formally analyzed by Hochreiter [15] and later by Bengio et al. [16].

LSTM networks, introduced by Hochreiter and Schmidhuber [17], addressed these limitations through their innovative cell state architecture. GRU networks, proposed by Cho et al. [18], later offered a simplified alternative to LSTM. CNNs, originally conceived by Kunihiko Fukushima as the Neocognitron [19], have evolved to become powerful tools for pattern recognition and feature extraction.

Recent studies have highlighted the importance of sensor data quality and processing in wearable applications. The integration of smartwatches into Internet of Things (IoT) frameworks, as discussed by Takiddeen and Zualkernan [7], presents both opportunities and challenges for real-time monitoring systems. However, as noted by

Lane et al. [9], deploying deep learning models on mobile and embedded devices remains challenging due to computational and power constraints.

III. METHODOLOGY

A. Data Collection and Processing

Initial training data was sourced from previous research [20][21], providing a foundation for model development. This was supplemented with new data collected using a Samsung Galaxy Watch 6 equipped with an LSM6DSO 6-axis IMU sensor. Following the methodology established by Windler et al. [22], a consistent sampling rate of 10 Hz is maintained across all data sources to ensure compatibility between training and deployment environments.

Signal quality was enhanced through multiple preprocessing steps:

- Nearest neighbor interpolation for consistent sampling, addressing the challenge of variable sensor sampling rates identified in [23]
- DC offset removal by subtracting the average value of each axis over time.
- Gaussian filtering for noise reduction, implemented using one dimensional gaussian filter provided within the SciPy python package, with default sigma values [24]
- Standard scaling for normalization, ensuring consistent feature ranges ($\mu = 0$; $\sigma = 1$) across different sensor axes.

To address demographic variations in movement patterns, we incorporated data gathered from [25] regarding the simulation of older adult movement patterns during data collection. This approach helps ensure the model's applicability across different age groups.

B. Signal Processing Techniques

To extract and enhance salient features from the raw sensor data, different Signal Processing techniques are applied, each of the following.

Principal Component Analysis (PCA) was implemented following the methodology described by Wold et al. [26], aiming to reduce dimensionality while retaining maximum variability within the data. This method is notably effective for handling correlated variables [27].

The Fourier Transform (FT) was implemented using the Fast Fourier Transform algorithm to leverage computational efficiency. FT enables frequency-domain analysis of periodic signals [28], making it especially suitable for the identification of repetitive activities such as walking.

Empirical Mode Decomposition (EMD) was executed according to the original procedure by Huang et al. [29]. EMD decomposes complex signals into Intrinsic Mode Functions (IMFs), which facilitates the analysis of non-linear and non-stationary signals [30].

The Hilbert-Huang Transform (HHT) integrates Empirical Mode Decomposition with the Hilbert spectral analysis, providing detailed time-frequency representation of signals [31]. This technique effectively captures dynamic and varying characteristics in signal behavior [31].

Each transformation method can be sequentially evaluated for its effectiveness in extracting meaningful features, improving classification accuracy, and maintaining computational efficiency, reflecting considerations critical due to resource limitations inherent in smartwatch deployments [9].

C. Neural Network Architectures

To benchmark model families under truly comparable conditions, we wrapped every network in an agent class that exposes the same fit-evaluate-save interface and inherits a common training configuration: 100-step sequences, batch size 64, Adam (lr = 1×10^{-3}), categorical cross-entropy, and early stopping with a patience of 10–20 epochs. The five agents differ only in the layers that transform the input stream.

- RNN agent: three SimpleRNN layers (256 → 512 → 256 units, tanh, 0.3 dropout) capture temporal context, followed by two dense layers (128 → 64, tanh) and a soft-max output.
- LSTM agent: identical topology but with LSTM cells (128 → 256 → 128 units) that retain long-range dependencies while mitigating vanishing gradients.
- GRU agent: a lighter three-layer GRU stack (64 → 128 → 64 units, 0.2 dropout) with dense layers (32 → 16, tanh), trading a smaller footprint for faster convergence.
- CNN agent: three Conv2D blocks (32, 64, 128 filters; 3 × 3 kernels; ReLU) each followed by 2 × 1 max-pooling compress the spectro-temporal representation; a 128-unit dense layer and soft-max complete the classifier.
- CNN ‑ LSTM agent: convolutional features are flattened via TimeDistributed and streamed into two LSTM layers (64 → 32 units, 0.3 dropout) before a 32-unit dense layer and soft-max. This hybrid marries local pattern extraction with sequence modelling.

Because all hyper-parameters outside the feature extractor are shared, performance differences can be attributed purely to the architectures themselves rather than to training-regime artefacts.

IV. RESULTS

A. Model Performance

All architectures except RNN achieved high accuracy When evaluated on raw inertial signals (Table 1), the convolutional (CNN), long short-term memory (LSTM), gated recurrent unit (GRU), and hybrid CNN–LSTM architectures all achieved near-perfect classification accuracies (0.9998–0.9999), whereas the vanilla recurrent network (RNN) yielded a markedly lower accuracy of 0.5747. Applying principal component analysis (PCA) produced only marginal improvement for the RNN, while elevating the CNN to perfect performance and slightly enhancing the hybrid model. Empirical Mode Decomposition (EMD) had the most uniformly positive effect on the RNN, boosting its accuracy to 0.9783, and it maintained or slightly improved the performance of all other models (CNN = 0.9999; LSTM/GRU = 1.0000; CNN–LSTM = 0.9998). The Hilbert–Huang Transform (HHT) exhibited a similar pattern: the RNN rose to 0.9617, the CNN slightly decreased to 0.9988. These results underscore that while empirical decompositions (EMD, HHT) effectively condition data for recurrent architectures, pure spectral filtering (Fourier) may inadvertently disrupt the feature hierarchies learned by convolutional and hybrid models (see Table 1).

TABLE I. MODEL PERFORMANCE

Transfor-mation	Tested model					
	Evaluation Metrics	*CNN*	*RNN*	*LSTM*	*GRU*	*CNN-LSTM*
Raw	Accuracy	0.9999	0.5747	0.9999	0.9999	0.9998
	Precision	0.9999	0.5255	0.9999	0.9999	0.9998
	Recall	0.9999	0.5747	0.9999	0.9999	0.9998
PCA	Accuracy	1.0000	0.6008	0.9999	0.9999	0.9999
	Precision	1.0000	0.5361	0.9999	0.9999	0.9999
	Recall	1.0000	0.6008	0.9999	0.9999	0.9999
Fourier	Accuracy	0.9661	0.5497	0.9977	0.9999	0.5497
	Precision	0.9490	0.3022	0.9977	0.9999	0.3022
	Recall	0.9661	0.5497	0.9977	0.9999	0.5497
EMD	Accuracy	0.9999	0.9783	1.0000	1.0000	0.9998
	Precision	0.9999	0.9781	1.0000	1.0000	0.9998
	Recall	0.9999	0.9783	1.0000	1.0000	0.9998
HHT	Accuracy	0.9988	0.9617	1.0000	1.0000	0.5497
	Precision	0.9988	0.9626	1.0000	1.0000	0.3022
	Recall	0.9988	0.9617	1.0000	1.0000	0.5497

a.

B. Smartwatch Application

An Android Wear application, developed in Kotlin, continuously acquires tri-axial accelerometer and gyroscope signals to enable on-device, real-time activity classification. As shown in Figure 1, the user interface displays a dynamic bar chart of model-predicted confidence scores and incorporates an opt-in toggle for asynchronous data streaming to a remote server. To meet the stringent CPU, memory, and power budgets of a smartwatch, we convert our neural network to a TensorFlow Lite (TFLite) format, achieving a significant reduction in binary size and inference latency without compromising classification accuracy. This architecture demonstrates that sophisticated convolutional–recurrent pipelines can be effectively deployed on resource-limited wearable platforms, paving the way for continuous, unobtrusive monitoring of activities of daily living.

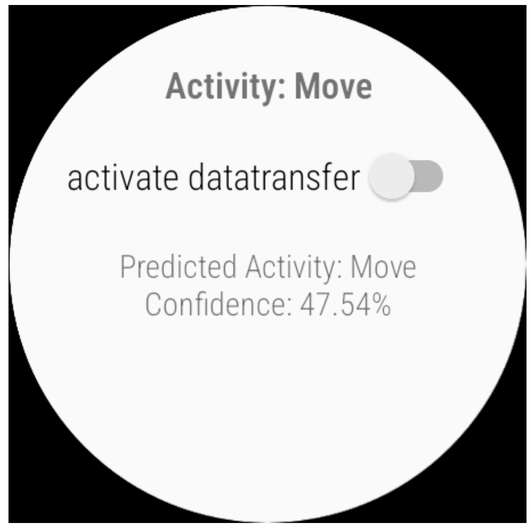

Figure 1. Smartwatch application interface for real-time activity recognition, displaying the predicted activity, confidence scores, and data transfer toggle.

C. Discussion

The deployment of the developed machine learning models on smartwatches highlighted several critical challenges that must be addressed to facilitate effective and continuous real-world use. Key challenges encountered during deployment included battery optimization, real-time processing constraints, the necessity for personalization, and variability in sensor data quality. Specifically, battery optimization emerged as a significant issue, as continuous model inference and sensor activity led to accelerated battery depletion, limiting the device's operational duration. Real-time processing constraints were observed due to the limited computational resources inherent to smartwatches, impacting the responsiveness and efficiency of the classification tasks. The need for personalization became apparent as performance variations were observed across different users and devices, highlighting that static, pre-trained models may not generalize well across diverse real-world conditions. Additionally, variable sensor data quality introduced inconsistencies, influencing the model's accuracy and reliability.

To overcome these limitations and enhance the deployment feasibility of activity classification models on wearable devices, several avenues for future research are recommended, as follows.

1) Development of efficient on-device learning mechanisms: Research should focus on implementing lightweight and computationally efficient on-device learning algorithms capable of continuous adaptation to individual user patterns, thereby enhancing personalization and mitigating performance degradation.
2) Battery consumption optimization: Further research is needed into advanced power management strategies, sensor management optimizations, and computational reductions (e.g., pruning, quantization) to extend battery life without compromising model accuracy.
3) Investigation of transfer learning approaches: Exploring transfer learning could facilitate more rapid personalization by leveraging pre-trained models adapted efficiently to new users with minimal data collection, addressing variability in user behavior and sensor conditions.
4) Integration with eldercare systems: Future studies should consider the integration of activity recognition systems with broader eldercare management platforms to improve the practicality and applicability of these models in monitoring daily activities, supporting elderly users, and enhancing their overall quality of life.

A pivotal evolutionary step to address the observed high variance in individual movement patterns would be to train user-specific models directly on the smartwatch. This approach would significantly enhance the adaptability and precision of activity recognition systems, thus improving their robustness and reliability in personalized, real-world scenarios.

V. Conclusion

This study confirms the viability of accurate human activity recognition using smartwatch sensor data and deep learning models. While advanced neural network architectures such as CNNs and LSTMs achieve high performance with minimal benefit from traditional signal processing techniques, these methods still hold value in enhancing simpler models or improving model efficiency. Importantly, the work underscores the practical constraints of deploying such models on resource-constrained wearable devices. Future research should prioritize energy-efficient inference, explore lightweight architectures, and investigate on-device learning strategies to enable adaptive, real-time activity recognition within the limited computational and power budgets of smartwatches.

References

[1] Statistisches Bundeamt [Statistical Bureau], “Germany’s Population by 2060”, Accessed: Sep. 23, 2024. [Online]. Available: www.destatis.de.

[2] Deutscher Pflegerat [German Nursing Council], “Alle Hebel zur Bewältigung der Pflegekrise umstellen | Deutscher Pflegerat [Change all levers to tackle the nursing crisis | German Nursing Council]” Accessed: Jun. 15, 2024. [Online]. Available: https://deutscher-pflegerat.de/profession-staerken/pressemitteilungen/loesung-der-pflegekrise-benoetigt-weit-mehr-als-350.000-pflegekraefte-innerhalb-der-naechsten-zehn-jahre

[3] WHO, “mHealth: new horizons for health through mobile technologies: second global survey on eHealth,” 2011, Accessed: Sep. 23, 2024. [Online]. Available: https://iris.who.int/handle/10665/44607

[4] M. Mubashir, L. Shao, and L. Seed, “A survey on fall detection: Principles and approaches,” 2012, doi: 10.1016/j.neucom.2011.09.037.

[5] S. Patel, H. Park, P. Bonato, L. Chan, and M. Rodgers, “A review of wearable sensors and systems with application in rehabilitation,” J Neuroeng Rehabil, vol. 9, p. 21, 2012, doi: 10.1186/1743-0003-9-21.

[6] S. G. G. Selen, N. A. A. Rahman, and K. S. Harun, “A study of wearable IoT device (smartwatch) advantages, vulnerabilities and protection,” AIP Conf Proc, vol. 2802, no. 1, Jan. 2024, doi: 10.1063/5.0183094/3126760.

[7] N. Takiddeen and I. Zualkernan, “Smartwatches as IoT edge devices: A framework and survey,” 2019 4th International Conference on Fog and Mobile Edge Computing, FMEC 2019, pp. 216–222, Jun. 2019, doi: 10.1109/FMEC.2019.8795338.

[8] F. Stradolini, E. Lavalle, G. De Micheli, P. Motto Ros, D. Demarchi, and S. Carrara, “Paradigm-Shifting Players for IoT: Smart-Watches for Intensive Care Monitoring”, 2017, doi: 10.1007/978-3-319-58877-3_9.

[9] N. D. Lane, S. Bhattacharya, A. Mathur, P. Georgiev, C. Forlivesi, and F. Kawsar, “Squeezing Deep Learning into Mobile and Embedded Devices,” Pervasive Computing, 2017, doi: 10.1109/ISCA.2016.11.

[10] Q. Yang et al., “10 Challenging Problems in Data Mining Research,” Int J Inf Technol Decis Mak, vol. 5, no. 4, pp. 597–604, 2006.

[11] R. Lutze and K. Waldhör, “Smartwatch based tumble recognition — A data mining model comparision study.” 2016, IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom 2016),) took place 14-17 September 2016 in Munich, Germany, J. J. P. C. Rodrigues, Ed. IEEE, Piscataway, NJ, 1–6.

[12] K. Waldhör and R. Baldauf, “Recognizing Drinking ADLs in Real Time using Smartwatches and Data Mining,” Proceedings of the RapidMiner Wisdom Europe, S. Fischer, I. Mierswa and G. Schäfer, Eds. Shaker, Aachen, 1–18, 2015.

[13] J. J. Hopfield, “Neural networks and physical systems with emergent collective computational abilities (associative memory/parallel processing/categorization/content-addressable memory/fail-soft devices),” Proc. NatL Acad. Sci. USA, vol. 79, pp. 2554–2558, 1982, Accessed: Sep. 21, 2024. [Online]. Available: https://www.pnas.org

[14] J. L. Elman, “Finding Structure in Time,” Cogn Sci, vol. 14, pp. 179–200, doi: 10.1207/s15516709cog1402_1.

[15] S. Hochreiter, “Untersuchungen zu dynamischen neuronalen Netzen [Investigations into dynamic neural networks]” Diploma, Technische Universität München, vol. 91, no. 1, p. 31, 1991.

[16] Y. Bengio, P. Simard, and P. Frasconi, “Learning Long-Term Dependencies with Gradient Descent is Difficult,” IEEE

Trans Neural Netw, vol. 5, no. 2, pp. 157–166, 1994, doi: 10.1109/72.279181.

[17] S. Hochreiter, "Long Short-term Memory," Neural Computation MIT-Press, 1997.

[18] K. Cho et al., "Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation," pp. 1724–1734, Oct. 2014, doi: 10.3115/v1/D14-1179.

[19] K. Fukushima, "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," Biol Cybern, vol. 36, no. 4, pp. 193–202, Apr. 1980, doi: 10.1007/BF00344251/METRICS.

[20] K. Waldhör and R. Lutze, "ADL Detection Challenge at IEEE Healthcom 2016 Conference - Wearables at work.," ADL Detection Challenge at IEEE Healthcom 2016 Conference - Wearables at work, 2016.

[21] F. Full, "Methoden zur Generierung und Klassifizierung von Sensordatan für die ADL-Erkennung," Aug. 2019.

[22] T. Windler, J. Ahmed Ghauri, M. Usman Syed, T. Belostotskaya, V. Chikukwa, and R. Rêgo Drumond, "End-to-End Motion Classification Using Smartwatch Sensor Data", 2020, doi: 10.5445/KSP/1000098011/12.

[23] F. Gu, M. H. Chung, M. Chignell, S. Valaee, B. Zhou, and X. Liu, "A Survey on Deep Learning for Human Activity Recognition," ACM Comput Surv, vol. 54, no. 8, Nov. 2022, doi: 10.1145/3472290.

[24] SciPy, "gaussian_filter1d — SciPy v1.14.1 Manual." Accessed: Oct. 19, 2024. [Online]. Available: https://docs.scipy.org/doc/scipy/reference/generated/scipy.ndimage.gaussian_filter1d.html

[25] K. Waldhör, "Maschinelles Lernen und Smartwatches zur Unterstützung eines selbstbestimmten Lebens älterer Personen [Machine learning and smartwatches to support independent living for older people]" in Künstliche Intelligenz in Wirtschaft & Gesellschaft: Auswirkungen, Herausforderungen & Handlungsempfehlungen [Artificial Intelligence in Business & Society: Impacts, Challenges & Recommendations for Action], T. and K. O. Buchkremer Rüdiger and Heupel, Ed., Wiesbaden: Springer Fachmedien Wiesbaden, 2020, pp. 347–367. doi: 10.1007/978-3-658-29550-9_19.

[26] S. Wold, K. Esbensen, and P. Geladi, "Principal Component Analysis". 1987, Chemometrics and Intelligent Laboratory Systems 2, 1-3, 37–52, Doi: 10.1016/0169-7439(87)80084-9.

[27] I. T. Jolliffe and J. Cadima, "Principal component analysis: a review and recent developments," Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, vol. 374, no. 2065, p. 20150202, Apr. 2016, doi: 10.1098/rsta.2015.0202.

[28] E. J. Haugstvedt, "On the Potential of Utilizing Laboratory-Scale Experimental Setup as Proxy For Real-Life Applications: Time Series Analysis and Prediction for Hole Cleaning," 2023.

[29] N. Huang et al., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, pp. 903–995, 1998, doi: 10.1098/rspa.1998.0193ï.

[30] G. Rilling, P. Flandrin, and P. Gonçalvès, "On empirical mode decomposition and its algorithms", 2025, Accessed: June 19, 2025. [Online]. Available: https://inria.hal.science/inria-00570628/en/.

[31] S. Kizhner, T. P. Flatley, N. E. Huang, K. Blank, E. Conwell, and D. Smith, "On the Hilbert-Huang Transform Data Processing System Development", Accessed: June 19, 2024. [Online]. Available: https://ieeexplore.ieee.org/document/1367979.